JOURNAL OF ECONOMIC SURVEYS

doi: 10.1111/joes.12202

CHILDREN AND THEIR PARENTS: A REVIEW OF FERTILITY AND CAUSALITY

Damian Clarke*

Department of Economics
Universidad de Santiago de Chile

Abstract. Childbearing decisions are not made in isolation. They are taken in concert with decisions regarding work, marriage, health investments and stocks, as well as many other observable and non-observable considerations. Drawing causal inferences regarding the effect of additional children on family outcomes is complicated by these endogenous factors. This paper lays out the issues involved in estimating the effect of additional child births on family outcomes, and the assumptions underlying the range of estimators and methodologies proposed in the economic literature. The common pitfalls of these estimators are discussed, as well as their potential to bias our interpretation of the effect additional births have on children and parents, both in the existing literature and in future work in the face of changing patterns of childbearing and child-rearing.

Keywords. Causality; Childbirth; Contraceptives; Fertility

1. Introduction

Human decisions regarding births, and how these decisions affect individual outcomes, are central to human welfare. They are also widely relevant to the world population. In 2014, 18.7 per every 1000 people had a child, which is the equivalent of 4.3 births per second (CIA, 2014). Over the course of her lifetime, the average woman in 2013 will have 2.46 births, down from 4.98 in 1960 (The World Bank, 2015). The importance of choice and control over the timing and number of children has been documented as early as *c*1800 BC, with discussions of a range of contraceptive methods included in the Kahun Gynaecological Papyrus (O'Dowd and Philipp, 1994). Estimates suggest that between then and today, contraceptive use has grown to reach global coverage of 60% of all married, fertile-aged women (Darroch, 2013). This paper examines the effect of childbirth and birth timing decisions on human outcomes.

Rather than focus on correlations between fertility and other outcomes, this paper is centred on the *causal* analysis of the effects of fertility. I discuss the theoretical and empirical considerations required to infer causality in a behaviour which cannot be manipulated directly in an experimental context. Given the interruptive nature of child birth on a large range of other life outcomes, any study of causal effects must isolate changes in fertility from corresponding changes in simultaneously determined, or dynamically dependent, outcomes. As a simple example, if women jointly choose to exit the labour market and have a child, any inference regarding the effect of fertility on her or her child's *other* outcomes must be independent of her labour market choice.

*Corresponding author contact email: damian.clarke@usach.cl; Tel: +56227180750.

Since the boom in fertility-related research in microeconomics in the mid-1970s[1], a range of methodologies have been proposed to permit inference in precisely these circumstances. These include both fully structural estimation strategies, as well as reduced-form methods such as the use of instrumental variables (IVs), combining difference-in-difference (DD) with IV, the use of quasi-experimental fertility shocks or trying to artificially construct a treatment and control group using family members or other matching methods. In the sections which follow, I describe these methodologies, the empirical results they have given rise to, and the identifying assumptions that motivate causality in each case. I also discuss common threats to inference, and what, if anything, these threats imply with regard to the existing estimates in the fertility literature.

Empirical considerations regarding the causal effects of fertility require the consideration of (at least) three questions: ‘Effects on what?’, ‘Effects at what margin?’ and ‘Effects of what? (timing or quantity)’. Frequently, these questions are split further, and examined for specific groups of women or children.

Regarding the first question, the theoretical and empirical microeconomic literature has hypothesized that marginal births may have causal implications for many individual-level outcomes. This includes effects on children: their health indicators, cognitive and non-cognitive achievements, long-term labour market outcomes, education inputs and social outcomes such as age at marriage and crime incidence; as well as effects on parents. Parental outcomes often considered are centred around labour market participation and returns, rates of education completion, marriage market outcomes and socioeconomic indicators such as welfare receipt.

Human fertility decisions exist at two (very different) margins. The choice of whether or not to have a child (the extensive margin), and, conditional on choosing to have any children, the decision regarding *how many* births to have (the intensive margin). The nature of links between fertility decisions and outcomes vary substantially when considering the intensive and the extensive margin. The consensus in the literature is that the causal effects are certainly not a linear function of births, with important non-linearities, and indeed non-monotonic relationships, described between fertility and (some) of the previously mentioned outcomes. In order to quantify effects at different margins, a range of estimation samples and methodologies need be employed.

Finally, causal effects of fertility are not independent of mother’s age at birth. Often, rather than estimating the effect of a marginal birth, we will be interested in determining the effect of child birth at a particular age (such as during adolescence). Considerations of these effects have important life-cycle implications for future investment decisions.

The study of fertility is common to a huge range of fields: social sciences, physical sciences, demography and medicine, and in many sub-fields within disciplines. As a result, any discussion of the state of the field must be necessarily pointed. This paper firmly focuses on the effects of fertility on other individual-level outcomes, and not the determinants of fertility[2], the macroeconomic effects (Enke, 1966, 1971), or discussions of the broader effects of population control policies (Rosenzweig and Wolpin, 1986; Miller, 2010). This paper, and indeed the literature on which it is based, largely focuses on the effect of a woman or family’s fertility decisions on the mother’s life outcomes and the outcomes of her children. Some, although relatively less, focus is paid to the effect on her partner (if existing and present).

2. Causality and Fertility

2.1 *A Framework*

We will consider an outcome Y_i, for each member i of a sample, $i \in \{1, \dots, N\}$, where Y denotes an outcome variable of interest, and the sample is drawn from a population of childbearing mothers (or, as discussed later, their children). We are interested in determining the effect of manipulations of fertility, which we denote F_i, on our outcome variable of interest Y_i. It is assumed that Y_i is a function of fertility, an unobserved variable U_i, and a series of other variables which are summarized as the error term ε_Y:

$$Y_i = f_Y(F_i, U_i, \varepsilon_Y) \tag{1}$$

Fertility is assumed to be a function of the unobserved U_i and stochastic ε_F:

$$F_i = f_F(U_i, \varepsilon_F) \tag{2}$$

and finally $U_i = f_U(\varepsilon_U)$. These error terms $\boldsymbol{\varepsilon}$ are assumed mutually independent. To fix ideas, we could consider an outcome variable Y_i as average years of education of i's children, F_i as completed fertility and U_i as unobserved positive health behaviours of the mother. By iterative substitution of the $\boldsymbol{\varepsilon}$ terms into (1) and (2), it becomes apparent (in the defined system of equations) that changes in fertility and health are unrelated to ε_Y, but that both average years of education and fertility are related to unobserved maternal health behaviours. In other words, F_i and U_i are not functions of ε_Y; however, Y_i and F_i are functions of ε_U.

In causal terms as per Haavelmo (1943, 1944) (and particularly, the recent exposition in Heckman and Pinto, 2015), we are interested in the change in Y resulting from the hypothetical manipulation of fertility F, while other elements of the system of equations $(U, \boldsymbol{\varepsilon})$ remain unchanged. We define b as a particular draw of F, and are thus interested in the causal effect of manipulating fertility from b to $b + 1$, which throughout this paper we will call β:

$$\beta \equiv E_{U_i,\varepsilon_Y}[Y_i(b_i + 1) - Y_i(b_i)] \tag{3}$$

For now we make no distinction between different values of b in (3). Generally, however, we will be interested in at least two separate situations. The first, comparing having any children to having no children (the extensive decision), while the second refers to having $b + 1$ children versus having b children for $b \in 1, \ldots, k$ (the intensive margin). We return to discussions of b for different parities when discussing empirical results in the sections which follow.

The hypothetical manipulation envisioned by Haavelmo is generally not feasible in real-world fertility decisions. While various small- and large-scale programs exist which have experimentally varied the cost of family planning, or family planning information (for example, the Matlab experiment in rural Bangladesh), direct manipulations of fertility itself are neither practical nor ethical. And in observational studies using data over space or time, the presence of factors similar to U considerably hinders the estimation of causal effects. In the sections which follow we return to this system of equations, and outline the existing techniques which aim to recover causal estimates despite the lack of explicit exogenous manipulation of F.

2.2 *Questions and Applications*

Discussions of causality can be entirely agnostic to the deeper questions of why effects are observed, or why we may want to quantify causal effects at all.[3] Nonetheless, these are precisely the reasons that research into these questions is undertaken. In the first instance, the estimation of causal effects can allow us to test well-specified models or hypotheses regarding behavioural or biological mechanisms underlying the effect of fertility choices. A correctly estimated parameter in the fertility literature is generally of interest given its relation to deeper behavioural or technological implications rather than as a curio in its own right. And in the second case, these behavioural or technological implications are extremely relevant for policies implicitly or explicitly designed to impact human well-being. I discuss these 'whys' in what remains of this section.

A range of theories exists to explain why fertility choices may affect other human outcomes. Microeconomic theories of fertility choice have a number of roots, including biological [for example, the Trivers and Willard, 1973 hypothesis regarding mother's health, environment and investment in offspring has been used to motivate economic applications (Almond and Edlund, 2007)], behavioural or technological. The earliest work such as Becker (1960) approached the problem firmly through the lens of consumer behaviour. Utility maximizing individuals (or families) were assumed to choose the number of children that they would like to have in the same way that they were assumed to decide on other consumption: based on relative prices (or shadow prices) and a budget constraint. In later work (Becker, 1965), the idea of a household production function was introduced, in which both monetary and time costs of production and consumption were considered to motivate household decisions.

These theories of 'demand for children' gave rise to theoretical predictions regarding household fertility decisions. Perhaps most central to these is the quantity–quality (QQ) trade-off, which posits that an inverse relationship will exist between child quality and child quantity. This relationship is suggested to exist given that children are a special composite 'good', where along with deciding how many are desired (an extensive decision), parents decide on how much to invest in child quality (an intensive decision). I return to discuss the QQ trade-off later in this paper.

These demand-based theories which can also incorporate time endowments of family members give rise to a number of other hypothesized relationships, including an inverse relationship between costs of childbearing (including outside labour market options for parents), and the number of children born. Concerned with the narrow nature of demand-based theories where all variation in fertility is due to parental tastes and the prices they face, Easterlin (1975) proposed a more comprehensive theory, in which demand for children, the supply of children (the theoretical quantity of children if parents do not contracept) and the cost of fertility regulation interact to determine completed fertility. Easterlin provides a deeper analysis of the machinery behind these decisions. He discusses the 'basic' and 'proximate' determinants of fertility, which include (respectively) socioeconomic conditions, modernisation variables, cultural factors and genetic factors; and exposure to intercourse, fecundability, duration of postpartum infecundability and the use of fertility control. This theory conserved the most important predictions from Becker and coauthors' models, while also opening new considerations regarding the determinants and effects of fertility. Perhaps, most importantly, these theories opened the door to external determinants of fertility such as fertility control technologies as explicit determinants which may have a direct effect on fertility outcomes.

The search for the theoretical and empirical implications of fertility choices and determinants reaches far beyond the academic literature. These effects are frequently cited in policy documents drafted by governments and international organizations when defining, classifying or justifying policy choices which have the potential to remarkably change fertility choices, and life courses, of affected individuals. There exist a range of proclamations and charters defining reproductive rights for individuals. As early as 1968, the United Nations Declaration of Human Rights incorporated that '[p]arents have a basic human right to determine freely and responsibly the number and the spacing of their children' (United Nations High Commissioner for Human Rights, 1968), a statement echoed in many subsequent proclamations, including the Proclamation from the Cairo Program of Action in 1994, and in the current World Health Organisation definition of reproductive rights.

Currently, the proportion of countries whose governments explicitly state that they are trying to alter population levels is very high. In 2013, of the 66 countries classified as 'high fertility' (greater than 3.2 births per woman), 90% of these have explicit policies in place to lower fertility rates (United Nations, 2014). In many cases, these policies' stated aim is to allow families to access desired contraceptive technologies[4]. However, in many others, it is well recognized that high fertility has direct implications for development, including effects on educational attainment, and mother and child health (United Nations, 2014). In the remainder of this paper, I turn to the underlying estimates which are (at least partially) used to justify expensive and wide-reaching policies of these types.

3. The Effects of Family Size on Children

In 1973, the *Journal of Political Economy* released a special issue dedicated to new economic approaches to fertility. Interest in determining the causal effect of total fertility on the outcomes of children in the household blossomed from the articles it contained. A common theme in a number of articles in this issue (Becker and Lewis, 1973; De Tray, 1973; Willis, 1973) concerns a family's decisions regarding fertility (the quantity of children) and investments in child human capital (the 'quality' of children). Abstracting from intra-household variations in child quality[5], each of the aforementioned articles demonstrates the theoretical existence of a QQ trade-off[6].

The QQ trade-off described in the above series of articles as well as in Becker and Tomes (1976, 1986) owes to the joint entry of quality and quantity in the household budget constraint. As the number of children enters in the shadow price of quality, and the quality of desired children enters the shadow price for quantity, decisions regarding fertility and quality cannot be made in isolation. Holding all else constant, increases in fertility increase the shadow price of quality, and increases in quality increase the shadow price of the marginal birth. This considerably complicates causal inference. What's more, as recognized in early articles by Ben-Porath and Welch (1972) and Ben-Porath (1976), quality decisions may *directly* feed back to quantity via child mortality.

3.1 *Theory, Empirics and Their Interaction*

While a theoretical result, a number of articles – both historically and more recently – provide empirical support for the QQ trade-off when explicitly starting from the theoretical constructs of Becker and Lewis (1973). These papers link a utility-maximizing child investment decision directly to their estimation strategy. A classic example is Rosenzweig and Wolpin (1980a)'s work in resolving the prevailing intractabilities in estimating parameters from Becker and Lewis (1973)'s QQ model. While most recognized as the introduction of the twin instrument to the economic literature, this is directly borne from the restrictions imposed by a utility maximizing family jointly optimizing fertility choices with child investment decisions.

Further recent work starts explicitly from a (theoretical) QQ framework, however, aims to loosen the assumptions of early QQ models. Aizer and Cunha (2012), Mogstad and Wiswall (2016) and Brinch *et al.* (2012) all loosen assumptions regarding: (a) the assumed homogeneity of children or (b) the assumed homogeneity of parental *response* to all children. All of these studies begin with a parental utility function which includes both child quality and child quantity as competing choices, and indeed, attach a precise structure to this utility. However, these utility functions extend earlier models in permitting heterogeneity over children, the effect of siblings by birth order, or the way that parents treat children with different endowments. As in Rosenzweig and Wolpin (1980a), the models proposed by Aizer and Cunha (2012) and Brinch *et al.* (2012) are linked directly to estimated specifications, bridging the QQ theory and empirical findings.

Although none of these papers are fully structural, they take seriously the decisions and restrictions faced by parents in optimizing their fertility choices which were proposed in the earliest QQ models. This can be pushed even further in structural papers where estimation strategies *directly* interact with parental QQ behaviour. Models of these types have been shown to have considerable explanatory power in modelling observed fertility and human capital outcomes, and more importantly, to interact with and explain behavioural responses to post-birth educational and labour market policies (Todd and Wolpin, 2006).

More generally, a range of papers in the *macro*-economic literature examine the effect of aggregate fertility on human capital attainment and growth. Papers in this field – which started with the theoretical work of Becker and Barro (1988), Barro and Becker (1989) and Becker *et al.* (1990) – link individual decisions over fertility and child investment to aggregate macro trends in fertility and human capital [see, for example, de la Croix and Doepke, 2003 and references therein. These are also among the central issues examined in *Family Economics Writ Large* (Greenwood *et al.*, 2016)]. While clearly linking the economic theory of fertility choices and human capital investment with empirical results, these papers consider national-level measurements, and as such are not a central part of this review. An interesting, brief and cross-cutting overview of much of the original Beckerian QQ theory and its importance for present work across economic fields is provided by Doepke (2015). However, even when not directly invoked in reduced-form papers to explain how empirical results owe to optimal family behaviour, QQ theory casts a long shadow over the ways these papers are set up and estimated. The hypothesized effect of fertility on child quality is built directly in to many reduced-form models. I turn to discuss this work now.

3.2 *Observational Data*

Given the aforementioned theoretical structure of the relationship between child quality and child quantity, it is apparent that estimating OLS (Ordinary Least Squares) on observational data will lead to consistent estimates of β only in very particular circumstances. To see this, we return to equation (1). If we consider standard OLS with a linear model, we re-write (1) as:

$$Y = \alpha + \beta F + U + \varepsilon_Y$$

where we assume that $E[\varepsilon_Y] = 0$. To estimate β from the above, we can consider conditioning on two distinct values of F:

$$\begin{aligned} E[\hat{\beta}] &= E[Y|F = b+1] - E[Y|F = b] \\ &= E[\beta(b+1) + U|F = b+1] - E[\beta(b) + U|F = b] \\ &= \beta + \{E[U|F = b+1] - E[U|F = b]\} \end{aligned} \quad (4)$$

Thus, (4) is only identical to the causal estimate in (3) in a very limited set of circumstances: above this is when $E[U|F = b+1] = E[U|F = b]$. This is simply a specific example of the well-known OLS requirement that the independent variable of interest (F) must be uncorrelated with the omitted error term, given that $\text{plim}(\hat{\beta}) = \beta + \text{Cov}(F, U)/\text{Var}(F)$. Where variation in fertility in a cross-sectional dataset is correlated with movements of other variables related to the outcome of interest (which here we summarize as U), we will fail to identify the true causal effect of fertility given the lack of Haavelmo's hypothetical manipulation of F.

Due to the limitations laid out in the preceding paragraph, very few papers in the literature aim to infer causality by estimating linear models with cross-sectional data. Early work, such as Desai (1995), provides cross-sectional descriptive evidence to document correlations, while Hanushek (1992) estimates some cross-sectional (though value-added) models. However, many papers which use alternative methods to infer causality (discussed in the sections which follow) estimate OLS as a base specification, which can provide some information on the type and degree of bias in OLS. Beyond recognizing that a bias is likely to exist, relatively few of these papers provide an explicit discussion of why this may be. Notable exceptions include Qian (2009), who suggests joint parental preferences for more education and fewer children as well as optimal stopping rules which depend on the quality of the first child, and Black *et al.* (2010), who additionally note that family size effects are confounded with birth order effects. Indeed, there are a number of reasons why one may be concerned that bias remains. These include parental education, discount rates, maternal health or network effects driving both fertility and child quality. Generally, it seems likely that these factors will induce a negative bias in OLS estimates of the effect of fertility, given that factors which lead to fewer births (contraceptive knowledge, opportunity cost of time, aspirations and so forth) also seem likely to drive greater investments in children who are eventually born. However, there is no reason why this theoretically must be the case, as fertility choices may interact in a range of ways with many different unobserved characteristics of mothers or families. Empirically, although it largely seems to be the case that OLS estimates of the effect of fertility are lower (more negative) than more credibly causal estimates, there are some cases, particularly in more developed countries, where this is not the case (and recent evidence from Myrskylä *et al.*, 2009 suggests that declines in fertility with development may be reversed at some point).

3.3 *Instrumental Variables*

In systems of equations of the type described in Section 3, one way to drive inference is through the use of shifters (or IVs) which affect the quantity of one of the variables without affecting the other. In order to identify the effect of fertility on children's outcomes, this IV must affect only fertility, with no indirect effects on quality.[7] Returning to the nomenclature introduced in Section 2, consider Y_i as child quality, F_i as child quantity and the unobserved U_i, all generated as described in Section 2. However, now consider the case where a new variable from outside the system is observed, denoted by Z_i. Z_i is assumed to directly affect F_i:

$$F_i = f_F(U_i, Z_i, \varepsilon_F)$$

The function 1 is unchanged, reflecting the fact that the only channel with which Z_i affects Y_i is through F_i (in other words, the exclusion restriction holds). Finally, to close, assume that $Z_i = f(\varepsilon_Z)$, and once again the error terms $\boldsymbol{\varepsilon}$ are assumed mutually independent.

The above situation leads to an explicit way of generating the hypothetical variation discussed in Section 2. By taking advantage of variation in F induced by variation in Z, the effect of F on Y can be identified in the absence of any movement in U. The most simple way to consider this is by observing the Wald estimator:

$$\hat{\beta} = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[F|Z=1] - E[F|Z=0]} \tag{5}$$

Here, rather than explicitly being based on a movement from $F = b$ to $F = b + 1$ estimation is driven by the effect which Z has on Y, scaled by the degree to which it moves F. If the instrument increases birth by exactly one, then (5) collapses to an expression similar to (4). Fundamentally, in strategies of this type, the identifying assumption shifts from concerns regarding correlations between U and F to correlations between U and Z. Consistent causal estimation now requires that $\text{Cov}(U, Z) = 0$.

The earliest discussion of these types of shifters and the corresponding exclusion restriction required for the estimation of the causal effects of fertility was in Rosenzweig and Wolpin (1980a). They point out that if multiple births are unanticipated, their occurrence will cause some families to exceed their desired fertility, shifting the total number of births in the absence of any change in parental considerations of quality investments. This has motivated estimation in a number of papers, where twin births are employed as IVs. Twin instruments have been employed in a range of contexts and to examine various different 'quality' outcome variables of children. These include Black *et al.* (2005), Cáceres-Delpiano (2006), Li *et al.* (2008), Dayioğlu *et al.* (2009), Sanhueza (2009), Black *et al.* (2010), Angrist *et al.* (2010), Fitzsimons and Malde (2010), Marteleto and de Souza (2012) and Ponczek and Souza (2012), and focus on child quality measures including years of education, IQ, private school enrolment, BMI and height, college completion and age at marriage. The evidence on the existence of a QQ trade-off in these studies is mixed, although recent influential results suggest that the evidence in favour of a trade-off may be weak. In Table 1, I lay out outcome variables, contexts and estimates of β presented in the IV literature.

As per the above series of equations, causal estimates rely on the fact that Z truly is independent of U. This has been questioned in a number of ways. Rosenzweig and Zhang (2009) suggest that the close birth-spacing of twins, and the fact that twins have lower health stocks at birth (Almond *et al.*,

Table 1. Empirical Results: Fertility and Child Outcomes (IV)

Author	Country	Outcome	Estimate(Std. Err.)
		Panel A: Twins	
Black *et al.* (2005)	Norway	Years of Educ	−0.16(0.44)
Cáceres-Delpiano (2006)	USA	Private School	−0.000(0.005)
		Behind cohort	0.005(0.004)
Li *et al.* (2008)	China	Educ (categorical)	−0.027(0.014)
		Educ (enrolment)	−0.025(0.013)
Dayioğlu *et al.* (2009)	Turkey	Attendance	0.203(0.245)
Sanhueza (2009)	Chile	Years of Educ	−0.280(0.092)
Black *et al.* (2010)	Norway	IQ (standardized 1-9)	−0.170(0.052)
		Complete Highschool	−0.039(0.015)
Angrist *et al.* (2010)	Israel	Years of Educ	0.167(0.117)
		Some college	0.059(0.036)
		College grad	0.052(0.032)
Fitzsimons and Malde (2010)	Mexico	Years of Educ (F)	0.096(0.063)
		Enrolment (F)	−0.019(0.014)
Ponczek and Souza (2012)	Brazil	Years of Educ (F)	−0.634(0.194)
		Years of Educ (M)	−0.060(0.164)
		Panel B: Gender Mix	
Black *et al.* (2005)	Norway	Years of Educ	0.280(0.060)
Conley and Glauber (2006)	USA	Private school	−0.061(0.021)
		Grade repetition	0.007(0.004)
Lee (2008)	Taiwan	Total ln(educ spend)	0.328(0.088)
Angrist *et al.* (2010)	Israel	Years of Educ	−0.067(0.120)
		Some college	−0.025(0.025)
		College grad	−0.032(0.022)
Becker *et al.* (2010)	Prussia	Enrolment	−0.430(0.189)
Black *et al.* (2010)	Norway	IQ (standardized 1-9)	0.065(0.074)
		Complete Highschool	−0.019(0.021)
Kumar and Kugler (2011)	India	Years of Educ	−0.363(0.061)
Fitzsimons and Malde (2014)	Mexico	Years of Educ (F)	−0.015(0.125)
Millimet and Wang (2011)	Indonesia	BMI for Age	0.049(0.013)
		Panel C: Fertility Shock	
Bougma *et al.* (2015)	Burkina Faso	Years of Educ	−0.99(0.40)
Maralani (2008)	Indonesia	Years of Educ (early)	−0.167(0.117)
		Years of Educ (late)	−0.054(0.055)

(*Continued*)

Table 1. *Continued*

Author	Country	Outcome	Estimate(Std. Err.)
		Panel C: Fertility Shock	
Hotz *et al.* (1997)	USA	Complete highschool	−0.147(0.406)
Dang and Rogers (2013)	Vietnam	Years of Educ	−0.589(0.392)
		Private tutoring	−0.318(0.147)

Note: In the case that various samples are reported in the papers, the pooled estimate for female and male children of all women from the most recent time period is reported. In the case of twins estimates, the '3+' sample (twins at third birth as an instrument for fertility in families with at least three births) is reported. Where the original studies report p-values associated with estimates rather than standard errors, these are converted into standard errors for inclusion in this table.

2005) means that parents may change behaviours to reinforce or compensate for intra-household human capital differences. If parents act to compensate positive health endowments, they will invest more in larger non-twin children, and estimates using twins will understate the magnitude of the trade-off (and vice-versa if parents compensate for poor health endowments). While this can be tested directly, it requires data on early life human capital endowments such as birthweight. Bhalotra and Clarke (2015) question the exogeneity assumption in another way. They demonstrate that healthier mothers are more likely to take twin births to term, and at the same time that healthier mothers are more likely to have additional resources to invest in child quality later in life. At the very least, however, both of these critiques will lead to predictable biases in estimates of β, resulting in bounds on the effect of fertility on child outcomes.

A frequently used alternative to twin births consists of instrumenting with the gender mix of children born in the family. Generally, it is argued that parents prefer to have offspring of both genders (Conley and Glauber, 2006; Angrist *et al.*, 2010; Becker *et al.*, 2010; Millimet and Wang, 2011; Fitzsimons and Malde, 2014), and so those having various children of the same sex are more likely to continue childbearing. Alternatively, in some circumstances, it is argued that parents have a son preference, and so are more likely to continue after having early-birth girls (Lee, 2008; Kumar and Kugler, 2011). In both cases, these are empirically shown to be important drivers of fertility. Again, like estimates driven by twin births, empirical results are mixed, although recent evidence seems to point to statistically insignificant (though nearly universally negative) estimates of the trade-off, as outlined in panel B of Table 1.

Causality in this case requires that child sex mix has no direct effect on quality. This implies (among other things) that there are no gender-specific economies of scale which facilitate child quality investments more when children are of the same sex (Butcher and Case, 1994). While one could argue (and indeed hope) that goods which could be employed in the household for boy's education could also be employed for girl's education, generally there are other concerns. Dahl and Moretti (2008) show that gender composition affects the likelihood that parents live together. Butcher and Case (1994) provide extensive discussion of the potential that different child gender mixes may affect child costs, and demonstrate that in the USA, girls with sisters are significantly less educated than girls with brothers, postulating that this may be due to a reference group effect where parents have lower aspirations for their children when all children are girls. Concerns such as these cast doubt on the validity of the exclusion restriction described earlier in this section.

A range of other instruments have been proposed, including infertility (Bougma *et al.*, 2015), miscarriage (Hotz *et al.*, 1997; Maralani, 2008; Miller, 2009) and distance to family planning (Dang and Rogers, 2013). The outcomes and empirical results related to these studies are displayed in Table 1. While these instruments – all generally related to the ability to conceive or control conception – clearly drive fertility, in each case the exclusion restriction is questionable. This is explicitly treated in Hotz *et al.* (1997) who motivate techniques to recover bounds on the estimate of the effect of fertility. At the very least, in each case, if unhealthy women are more likely than healthy women to be infertile or suffer miscarriage, this suggests a positive bias in IV estimates of β.

Beyond general threats to inference discussed in this section, IV estimates lead to the question of 'inference for whom'. Estimates based on IV lead to a local average treatment effect (LATE), not an average treatment effect for the population in general (Imbens and Angrist, 1994; Ebenstein, 2009). This LATE implies that any estimates of β holds for that group of the population who would be induced to change their behaviour (i.e. their fertility) by the instrument in question. Thus, all instrumental estimates (even assuming causality) should be cast in terms of the sub-population (compliers) of interest. This is a point explicitly discussed in Angrist *et al.* (2010) who suggest that the twin instrument is relevant for the whole population, while sex-composition instruments are relevant for only certain groups. Rosenzweig and Wolpin (1980a)'s original article, although based on a reduced-form equation, suggests that twin births are relevant for a more specific group than that suggested by Angrist *et al.* (2010): namely those families who have a twin birth where the twin birth causes them to exceed their desired fertility.

3.4 *Natural Experiments*

An alternative manner to deal with correlation between F and U consists of taking advantage of externally defined (to U) reforms. If reforms are applicable to a sub-group of a particular population and are designed to affect fertility, this suggests a natural 'treatment' and 'control' group which can be compared. Those who receive coverage from the fertility reform are considered treated, and those who do not are considered as controls. If reforms are truly put in place for reasons entirely divorced from U, causal conclusions can be drawn regarding the effect of the reform. Typically, the effect of reforms is estimated using DD estimators. This compares pre-reform differences between treated and control units with post-reform differences, inferring that any change in the level of differences is driven by the reform, or stated in another way, that *no* differential and simultaneously occurring phenomena separate treatments from controls. This is the well-known 'parallel trends assumption' and is central to this line of inference.

These studies can be broadly split into two groups: those which examine the effect of public policies or other natural experiments on fertility itself, and those which leverage the externally defined effect on fertility to quantify the effect of fertility on some other outcome of interest. In the latter case, the differentiation between DDs and IV estimates is artificial, as the (DD estimated) effect of the policy on fertility is simply plugged in as the first stage in a 2SLS IV framework[8]. The first set of studies is of fundamental importance in analysing the *determinants* of fertility and the effect of new contraceptive methods on life-cycle childbearing, but do not directly quantify the causal effects of fertility itself. Nevertheless, given their relevance both as a first stage in causal estimates and as a reduced-form estimate itself, I outline a number of these studies in Table 2, before moving on to a more comprehensive discussion of their link to causal estimates.

Historically, many fertility reforms have been atypical when compared to other large publicly defined policies. The nature of fertility control technologies has meant that large changes in contraceptive availability have often occurred which were quite different (in both timing and design) to the stated aims of public planners. For example, the advent of the contraceptive pill in the 1950s, as well as the repeal of contraceptive laws in *Griswold v. Connecticut* in 1965 (Bailey, 2013) meant that the largest contraceptive reform in the USA in the 20th century occurred without an explicit public reproductive policy[9]. The piecemeal and unexpected nature of these reforms makes it difficult in many cases to quantify the effect of reforms on their stated aims due to the lack of centralized planning. Nevertheless, in some circumstances, comprehensive reviews can be conducted, focusing on the total effect of reproductive policy reforms on its actual aim. For example, Molyneux and Gertler (2000) analyse the effect of a national policy (Indonesia), while many analyses exist of the local experimental policy in Matlab Bangladesh (Joshi

Table 2. The Estimated Effect of Reforms on Fertility (Selected Studies)

Author	Abortion Effect	Pill Effect	Note
Angrist and Evans (1996)	−0.012(0.004)		a, ($x = 19$)
Levine *et al.* (1996b)	−0.019(0.007)		b
Gruber *et al.* (1999)	−0.059(0.005)		c
Bailey (2006)	−0.093(0.043)	−0.074(0.057)	a, ($x = 22$)
Guldi (2008)	−0.100(0.054)	−0.085(0.041)	
Bailey (2009)	−0.012(0.007)	0.028(0.048)	a, ($x = 22$)
Pop-Eleches (2010)	−0.068(0.012)		d
Ananat and Hungerman (2012)	−0.043(0.015)	−0.088(0.023)	

Note: All figures report the results of short-term access of a fertility reform on birth rates of young women unless otherwise specified in notes.
[a] Binary model with outcome 1 = first birth by age x. Bailey (2009) is an erratum for 2006.
[b] Estimate expressed as births per woman. Mean rate is 0.110.
[c] Estimate for states adopting 1974–1975. Estimate for 1971–1973 is −0.021(0.005).
[d] Estimate is for all women with primary or lower education aged 15 and over. For women with greater than primary education, the effect was slightly lower.

and Schultz, 2013). However, even in the absence of large-scale analyses of the effects of reproductive policy reforms on their stated aims, historical contraceptive reforms are still an excellent candidate to isolate the effect of changes in fertility on other outcomes of interest given that these events have large effects on fertility, and are potentially divorced from other simultaneous reforms or differential trends.

Of the large number of studies which use reforms of fertility-control policies[10] to examine the effect on fertility in a DD-style framework, only a relatively small number then employ this as the first stage to estimate the causal effect of fertility – the focus of this paper. Among those that *do* directly estimate the effect of fertility on child outcomes are Gruber *et al.* (1999), Ananat *et al.* (2009) and Ananat and Hungerman (2012). Gruber *et al.* (1999) examine the effect of fertility (via 2SLS) on the likelihood that a child lives with single parents, lives in poverty, receives welfare and on rates of infant mortality and low birth rates. Of these, it is suggested that fertility significantly increases the probability of living in poverty and having single parents, as well rates of infant mortality. Ananat *et al.* (2009) also examine these outcomes, and suggest that in the long run, the marginal child is more likely to have lived with a single parent, receive welfare and not have graduated college. Finally, Ananat and Hungerman (2012) return to these same outcomes and report a Wald ratio as in (5). These Wald estimates allow them to look at the characteristics of marginal child not born due to both the diffusion of the pill, and the legalization of abortion. Their results suggest that the two fertility control policies had remarkably different effects on marginal child characteristics. In agreement with the above studies, they suggest that the marginal child not born due to abortion legalization would have been 49.2% (se = 25.5) more likely to live in a welfare-receiving household. However, the marginal child not born due to pill diffusion looks very different: 8.0% (se = 4.4) *less* likely to belong to a welfare receiving household. These comparisons make manifestly clear the distinction between compliers for different instruments discussed at the end of Section 3.3. Given that the group of 'compliers' in the two policies had very different characteristics, estimated effects of fertility on outcomes are very different despite being plausibly causal in both cases.

Despite not directly estimating the causal effect of fertility on child outcomes, a number of other contraceptive-based natural experiment papers estimate the effect of the natural experiment on child outcomes. This reduced-form technique provides an estimate of the numerator of the ratio in (5), and so can be thought of as an unscaled estimate of the effect of fertility. Papers of this type include Pop-Eleches (2006) who finds that the outlawing of abortion in Romania worsened child education and labour market outcomes (conditional on parental characteristics), Bailey (2013) who reports that US contraceptive pill laws had long-standing impacts on children's eventual college completion, labour force participation and family incomes, and Bailey *et al.* (2016) who find that increases in contraceptive spending flowing from the War on Poverty and Title X had important effects in reducing the number of children living in poverty, and increasing average household incomes.

The validity of using policies of this type to isolate the effects of childbearing on child outcomes hinges upon the fact that the timing (or allowance) of fertility control reforms should not depend upon pre-existing differences between areas affected and those not affected by the reform. Any phenomena which will imply that 'treated' and 'untreated' areas would follow different paths *in the absence* of the reform will lead to inconsistent estimates of the effect of fertility on child outcomes. Generally, papers which propose estimation by leveraging reforms of this type run a series of tests, including event-study analysis, placebo regressions or a regression of receipt of treatment on pre-existing characteristics. For example, Bailey (2006) demonstrates that early access to the pill was unrelated to education, fertility norms, poverty rates, availability of household technologies such as washers and dryers, as well as labour market participation at a state level. The probability of early access is, however, related to the percent of Catholic residents in a state. However, directly testing the validity of such estimation methodologies is, of course, impossible, given that the counterfactual outcome – the world where the pill was not available – is never observed. This has led to back-and-forth discussion, questioning the validity of the use of policy-defined reforms to drive estimation (for example, see Joyce, 2013, who questions the exclusion restriction vs. Bailey *et al.*, 2013 who defend current state-of-the-art results).

While concerns that reforms may be systematically correlated with other unobservable factors are, of course, justifiable, the best sets of studies aim to use judiciously chosen control groups (including using women of different ages subject to the same geographic factors and institutions), to minimize concerns such as these. An alternative concern in identification strategies of this type surrounds the possibility that *local* reforms have more widely spread effects. For example, the availability of abortion in one region does not necessarily imply that nearby non-treated individuals cannot travel to treated areas, defying their quasi-experimental status to receive treatment (Levine *et al.*, 1999). Fortunately, violations of this type will, at worst, bias downwards estimated results. While this is reassuring if our object of interest is a bounds estimate, generally, with policy reforms, this will not be the case. Given the large cost that large contraceptive policies entail, it is important to be able to identify the precise effect of competing options. As a result, concerns such as these are often examined empirically, as is the case in Christensen (2012) and Bentancor and Clarke (Forthcoming).

Finally, a number of other natural experiments have been used in the literature to examine the effect of fertility on child outcomes[11]. Perhaps, most notably among these, Qian (2009) uses the relaxation of China's one child policy to estimate the causal effect of movements from one-child to two-child households. This study is unique for two reasons: the low parity shift of the experiment (an expansion from one to two children), and the fact that it finds that higher fertility in this case *increases* child schooling outcomes, especially among households who have two children of the same gender. These results suggest that estimates of fertility at the intensive margin may not be linear, and indeed may not even be monotonic by parity, changing from positive to negative at higher orders.

4. The Effects of Child Birth on Mothers

Beyond the analysis of a child's effect on his or her siblings' outcomes, a birth, at the extensive or the intensive margin, has myriad impacts on parents or other carers. The analysis of these effects has received considerable and ongoing attention in the economics literature. Much of the focus of this work falls on the effect of marginal births on mothers' labour market outcomes and trajectories.

Fleisher and Rhodes (1979) provide a summary of the early literature, with considerable coverage also provided in the *JPE* Fertility issue described in Section 3 (Gronau, 1973; Willis, 1973). As is the case with child investment and fertility decisions, choices regarding fertility, labour market participation and (adult) human capital attainment are linked, and dynamic in nature. Total fertility, and, if childbearing, birth timing have important impacts on labour market participation, non-labour market work, accrued experience and wages, while participation, experience and wages also influence timing and fertility decisions. Inferring causality in systems of this type is once again challenging, relying on the use of plausible instruments, natural experiments, structural estimation or a combination of methods.

The theoretical results of Becker (1965), Willis (1973) and Gronau (1973) link additional births with the availability of time – particularly of mothers – within the household. A range of structural papers interact directly with this theory, and bring its implications to estimable equations related to fertility and maternal labour market decisions. Seminal papers from Heckman and Willis (1976) and Hotz and Miller (1988) and more recent work such as Francesconi (2002) document the important dual nature of decisions relating to childbearing and the labour market. These structural papers make very clear the endogeneity in choices of labour supply and childbearing/child-rearing decisions. While the structural approach seeks to directly model the joint optimization of parents and families explicit in economic theory, a second, reduced-form approach seeks to isolate only the effect of fertility on labour market outcomes. This requires more fortuitous identifying information, and is what I turn to discuss in the remainder of this section.

4.1 *Natural Experiments*

Frequently, natural experiments of the type discussed in Section 3.4 are leveraged to quantify the effect of fertility on parent outcomes. Estimation is based on the fact that – at the level of the family – living in treatment or non-treatment areas is a randomly assigned variable which can be used to isolate effects on fertility in the absence of changes in other outcomes. This requires that these experiments be clearly demarcated, unexpected and not propagate from treated to untreated areas.

One of the most common natural experiments employed in these types of analyses is the arrival of new birth control technologies to a particular geographic area. There are a very large range of microeconomic studies which discuss the effect of these types of programs on a mother's total fertility. These can be broadly split into those which examine the short-run effects of contraceptives on fertility[12], and long-run analyses, which account for both short-run delays and long-run rearrangements in timing afforded by new technologies. Short-run analyses include those examining the contraceptive pill (Bailey, 2006, 2009; Christensen, 2012), abortion (Levine *et al.*, 1999; Guldi, 2008), the morning after pill (Durrance, 2013; Gross *et al.*, 2014; Bentancor and Clarke, Forthcoming) and medicare access (Kearney and Levine, 2009), while those examining the long-run effects of contraceptive reform on completed fertility include (among others) Bailey (2010, 2013, 2012) for the contraceptive pill and Angrist and Evans (1996) and Ananat and Hungerman (2012) for abortion.

Once again, however, beyond the direct relevance of this swath of studies for policy focused on fertility control, in order to apply these results to *causal* analysis of parental outcomes, the specifications discussed above can only act as a first-stage effect. For the full system of equations, we are interested in a two-step process: first, quantifying the effect of reforms on fertility, and then, from this, the flow-on effect that exogenous shifts in fertility have on mother (or carer) outcomes.

Only a subset of papers which focus on fertility reforms then go on to examine the second stage of interest here. As discussed in Sections 3.3 and 3.4, consistent estimation relies on an exclusion restriction assumption, whereby the only effect of the program on outcomes is driven by its effect on fertility. Ananat and Hungerman (2012) use both the pill and abortion to examine different groups of compliers, and report Wald estimates of the effect of fertility on single parenthood: for pill-compliers, marginal fertility reductions occur in contexts with less single parenthood, while the reverse is true for abortion compliers. Angrist and Evans (1996) report similar estimates for 1970 abortion reforms in the USA. They report that the effects of a particular type of childbearing (teen and unmarried) reduces the education and employment probability, particularly of black women. Other significant outcomes discussed in this framework include

Bailey *et al.* (2012) and Bailey (2006, 2013), who show that it has effects on wages over the life cycle or female labour force participation rates, and Christensen (2012) who finds (reduced form) effects on cohabitation. Interestingly, the work of Edlund and Machado (2015) shows how simultaneous natural experiments can produce unexpected interactions. Namely, they document that the invention of the oral contraceptive pill and allowances for access to young married – but not young unmarried – women resulted in increases in early marriage paired with longer fertility delay due to pill access, and *higher* human capital accumulation among this group of women.

4.2 *Instrumental Variables*

The use of IVs to examine the effect of fertility on *mothers*' outcomes (rather than children's outcomes as described in Section 3.3) follows a similar logic to that outlined in equation (5). An external variable which has strong effects on fertility but no direct effects on the outcome of interest *except* via its effect on fertility can be used to drive causal estimates. IV estimates are a popular methodology employed to determine the effect of fertility on mothers.

Outcome variables of interest are typically related to parental labour force outcomes[13], including female labour force participation (Agüero and Marks, 2008, 2011; Chun and Oh, 2002; Cáceres-Delpiano, 2008; Angrist and Evans, 1998), or earnings (Hotz *et al.*, 1997, Jacobsen *et al.*, 1999, Cáceres-Delpiano, 2006). A range of instruments has been proposed including twins, as in Section 3.3 (Rosenzweig and Wolpin, 1980b; Bronars and Grogger, 1994; Jacobsen *et al.*, 1999; Cáceres-Delpiano, 2012), gender mix (Agüero and Marks, 2008, 2011; Chun and Oh, 2002) and fertility shocks (Rosenzweig and Schultz, 1987; Cristia, 2008; Miller, 2011)[14]. The signs and magnitude of existing estimates of the effect of fertility on labour market outcomes largely point towards significant negative impacts, though not universally so. A summary of point estimates and confidence intervals of estimates is presented in Table 3.

These point estimates appear to be largely negative, with only two (from the same context and cohorts) suggesting non-negative results of fertility on female labour force participation or hours worked. What is more, these are largely statistically significant negative estimates, despite the well-known caveat that IV estimates typically suffer from very wide confidence intervals (Angrist *et al.*, 2010). However, as discussed earlier, all of these estimates are LATEs which hold for particular populations and compliers, and so do not provide external validity for inference in other populations. However, a literature pointing in the direction of a negative effect is suggestive that this result could be observed in other contexts. This is something examined extensively by Deheija *et al.* (2015) who, perhaps unsurprisingly, find that quasi-experimental evidence generalizes more readily to countries which share closer geographical, education, time and labour force participation characteristics.

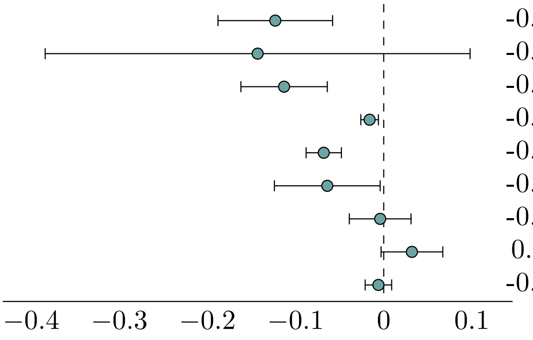
While the majority of these instruments can only be used to estimate the effect of fertility at the intensive margin, interestingly, those based on fertility shocks can also be applied to quantify the effects of extensive margin births. Cristia (2008), for example, proposes using the outcome of fertility treatments (pregnant or not) as an instrument, suggesting that delays in childbearing lead to an increase in wages and hours worked. Similarly, Lundborg *et al.* (2014) report considerable and long-lasting effects of extensive margin fertility on hourly earnings, using IVF success to drive estimates.

Finally, miscarriage has been proposed as an alternative IV that can be employed to estimate the effect of child birth on maternal outcomes (Hotz *et al.*, 2005; Fletcher, 2012). This line of argument relies on fetal deaths in utero being randomly assigned to mothers, in order to compare treated (live births) to control (no live births) women. This also relies on miscarriage not having any other effect on (prospective) mothers' outcomes of interest, beyond its direct effect on fertility. Hotz *et al.* (2005) suggest that following this line of argument, early (teenage) childbearing is associated with small effects on educational attainment, and life cycle changes in labour market rates.

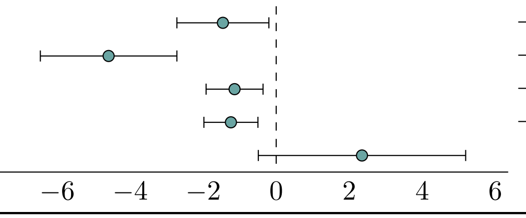
The use of this instrument is, of course, complicated if characteristics which predict miscarriage are also correlated with mother unobservables. Given that miscarriage is considerably more likely for

Table 3. Fertility and Mother's Labour Market Outcomes

Authors	$\hat{\beta} \pm 1.96\text{se}(\hat{\beta})$	$\hat{\beta}$
Labour Force Participation		
Bronars and Grogger (1994)		-0.123
Angrist and Evans (1996)		-0.143
Angrist and Evans (1998)		-0.113
Jacobsen *et al.* (1999)		-0.016
Cáceres-Delpiano (2006)		-0.068
Cáceres-Delpiano (2008)		-0.064
Agüero and Marks (2008)		-0.004
Angrist *et al.* (2010)		0.032
Agüero and Marks (2011)		-0.006



Authors	$\hat{\beta} \pm 1.96\text{se}(\hat{\beta})$	$\hat{\beta}$
Hours per Week		
Hotz *et al.* (1997)		−1.46
Angrist and Evans (1998)		−4.59
Jacobsen *et al.* (1999)		−1.14
Cáceres-Delpiano (2006)		−1.24
Angrist *et al.* (2010)		2.35



NOTES TO TABLE: Points represent coefficients, while error bars represent 95% confidence intervals. Estimates are ordered by date of publication. In the case that various samples are reported in the papers, the pooled estimate for all women from the most recent time period is reported. In the case of twins estimates, the 3+ sample (twins at third birth as an instrument) is reported.

unhealthy mothers, this seems likely, and a range of studies address these concerns. Foremost is Hotz *et al.* (1997) who discuss how to bound the effect of fertility where the IV is composed of a mixture of both women who randomly miscarry, and those who non-randomly miscarry. They show that tight bounds on the effect of fertility can be estimated, if the proportion of non-random and random miscarriages can be estimated. Applying these bounds estimates, they suggest that teenage childbearing significantly increases the number of hours worked during early adulthood, and (weakly) decreases the likelihood of completing a GED certificate in the USA. Fletcher and Wolfe (2009) provide additional discussion of the challenges in estimating causal effects using miscarriage. They suggest that unobserved *community*-level characteristics are likely correlated with miscarriage, and once including community fixed-effects find that teen childbearing reduces education and wages, and increase the likelihood of welfare receipt.

4.3 *Other Methods*

A range of other methods have been employed in the economic and non-economic literature to examine the effects of fertility on mother's outcomes. These involve RCTs (DiCenso *et al.*, 2002) between-effects using siblings (Geronimus and Korenman, 1992; Ribar, 1999; Holmlund, 2005), and other matching methods (Chevalier and Viitanen, 2003; Levine and Painter, 2003). In the case of the last two methods (between-effects and matching), the identification of casual effects relies on the comparison method fully controlling for relevant differences between those having children, and those not having children. In matching this collapses to an assumption regarding 'selection on observables' (which is to say that any characteristic predicting childbearing is observed by the econometrician), and in siblings or relative fixed effects, that, on average, those who become pregnant early in life are otherwise identical to those who become pregnant later in life. Ribar (1999) and Rosenzweig and Schultz (1985) provide additional discussion, and examination of the validity of these estimation techniques.

Finally, Rosenzweig and Wolpin (1980b) – in their initial proposition of twins as an exclusion restriction – return to the Beckerian (1973) simultaneous equation framework for fertility, child quality *and* life-cycle (mother's) labour supply. They are the first to use twins to estimate the structural equation linking fertility and labour supply. They estimate that for younger women, additional births reduce labour supply, but this fades as women age. Once again – as they indeed highlight – consistent estimation relies on twins being entirely orthogonal to labour supply. This assumption is questioned in previous sections of this paper. Using the presumed exogeneity of twins as an identifying assumption, Rosenzweig and Wolpin (1980b) provide a very interesting series of tests casting considerable doubt on the assumption that fertility is exogenous to labour supply decisions, as maintained in the prevailing literature at the time of their work.

5. Conclusion

This paper serves to provide an overview of the issues involved in the causal estimation of the effect of fertility on household outcomes. It surveys the wide range of methodologies employed in the existing economic literature, and discusses how various techniques aim to skirt issues of endogenous fertility choices. In each case, I outline the identifying assumptions implicitly or explicitly invoked, as well as the threats to which these are subject.

The evidence discussed in this paper is mixed. While there seems to be quite clear evidence in favour of moderate-to-large effects of marginal child births and early births on parental labour market outcomes, the existing micro-econometric child-level estimates are less compelling. Despite a large body of theoretical microeconomic work which posits that such a QQ trade-off may exist, causal estimates are certainly not conclusive, and seem to suggest that the trade-off is small or non-existent. While there are a number of papers which *do* find significant effects on a number of outcomes, these are context- and complier-specific.

These questions are of fundamental interest to future work in this field. While this topic has a long and rich history, its future is arguably just as rich. Changing patterns of childbearing mean that children are now born later, at lower parities, and with increasing planning and use of assisted reproductive technologies in many parts of the world. Similarly, advances in and diffusion of life saving technologies for mothers and children will drive broad changes in fertility patterns in other parts of the world. Given the importance of mortality reductions as an engine for declines in fertility, as developing countries continue along the demographic transition, even small direct effects of fertility on individual and family outcomes will have large aggregate effects. The interpretation of how these broad demographic shifts affect family outcomes will require an extension of the existing estimates and techniques to new life circumstances and situations.

Acknowledgements

I thank Sonia Bhalotra, Paul Devereux, James Fenske and Margaret Stevens for their many useful comments and extended discussions.

Notes

1. From 1970 onwards, the frequency of the occurrence of the word 'fertility' in titles of all articles published in the *Journal of Political Economy* is 46.5 per 100,000 words. Prior to the 1970s, the frequency was 1.45 per 100,000 words. Though admittedly based on a small sample, the frequency by decade is 11.17 ('60s), 72.72 ('70s), 66.33 ('80s), 29.41 ('90s) and 41.07 ('00s). It was not mentioned in titles of articles published between 1863 and the 1960s.
2. More recently, Kearney and Levine (2012) provide an extensive discussion of teenage childbearing and its determinants in the USA, Bailey (2013) reviews the old and new evidence on the effect of access to contraception, and Moffitt (2005) has provided a discussion and application of causal inference to the effects of teenage childbearing on child outcomes. All provide extremely useful reviews of these particular areas of the literature. The handbook chapter of Schultz (2008) is perhaps the definitive reference for micro-economists interested in an analysis with a very broad scope. There are many papers discussing education and fertility (i.e. Black *et al.*, 2008), which will not be discussed in this paper.
3. Discussions of how one should estimate causal parameters, and the limits of estimation without theoretical underpinnings are a topic of great debate, and have been for many years. As early as 1947 in 'Measurement Without Theory', Koopmans states:

   > for the purpose of systematic and large scale observation of such a many-sided phenomenon, theoretical preconceptions about its nature cannot be dispensed with, and the authors do so only to the detriment of the analysis.

   In discussions of parameter estimates, I do not zoom out to examine these deeper issues. Many resources discussing inference in economics provide fascinating insight into these issues, such as Keane (2010), Wolpin (2013) and references therein.
4. Recent estimates (Bongaarts and Sinding, 2011) suggest that approximately 40% of pregnancies in the developing world are unintended.
5. An extensive literature exists which looks at *intra*-household endowment and investment decisions among siblings. Behrman *et al.* (1982) provide initial discussion, and Aizer and Cunha (2012) embed these considerations in a QQ-type framework.
6. As Willis (1973) succinctly describes:

   > 'Thus, parents not only balance the satisfactions they receive from their children against those received from all other sources not related to children . . . , but they must also decide whether to augment their satisfaction from children at the 'extensive' margin by having another child or at the 'intensive' margin by adding to the quality of a given number of children.'

7. *Ie* the exclusion restriction must hold, implying that the estimation of the structural equation which contains quality on fertility and the instrument must result in a coefficient on the IV which is precisely equal to zero.
8. Duflo (2001) is a well-known example of this design. We discuss examples of this framework applied to fertility later in this section and in Section 4.1.
9. More recently, the US Supreme Court finding in *Burwell v. Hobby Lobby* affected birth control access in the absence of any reproductive policy. Similar examples exist in many other contexts.
10. For abortion, Ananat *et al.* (2007, 2009), Angrist and Evans (1996), Charles and Melvin (2006), Cook *et al.* (1999), Currie *et al.* (1996), Gruber *et al.* (1999), Guldi (2008), Kane and Staiger (1996), Levine *et al.* (1996b,a, 1999), Pop-Eleches (2010) and Pop-Eleches (2006); for the oral contraceptive pill: Ananat and Hungerman (2012), Bailey (2006, 2010, 2012, 2013), Christensen (2012), Goldin (2006), Goldin and Katz (2002) and Kearney and Levine (2009) and for the emergency contraceptive pill: Durrance (2013), Gross *et al.* (2014) and Bentancor and Clarke (Forthcoming).

11. Similarly, Bleakley and Lange (2009) use a natural experiment: the eradication of hookworm in USA, to test the QQ hypothesis. However, the elimination of hookworm is used as a shifter for child quality, *not* child quantity. This allows them to quantify the effect of quality increases on subsequent fertility decisions of households, and they find that increases in quality do lead to fertility declines in line with the QQ model discussed earlier.
12. This does not imply using data over a short time frame, but rather examining the effect of a birth control method up to an age *less* than the end of the fertile life (e.g. Bailey, 2006's focus on childbearing before the age of 22).
13. Largely, these papers focus on maternal labour market participation rates. Kim and Aassve (2006) study both mothers' and fathers' responses to fertility, using fecundity (births per attempt) as an instrument. They find that on average mothers reduce hours of work in the short run, while paternal hours of work increase (in rural areas).
14. Ribar (1994) proposes three alternative exclusion restrictions (age at first period, availability of Ob/Gyn and local abortion rates) for use in selection models. While the identification methodology is different to IV, the requirements for inferring causality are identical.

References

Agüero, J.M. and Marks, M.S. (2008) Motherhood and female labor force participation: Evidence from infertility shocks. *The American Economic Review* 98(2): 500–504.

Agüero, J.M. and Marks, M.S. (2011) Motherhood and female labor supply in the developing world: Evidence from infertility shocks. *Journal of Human Resources* 46(4): 800–826.

Aizer, A. and Cunha, F. (2012) The production of human capital: Endowments, investments and fertility. NBER Working Papers 18429, National Bureau of Economic Research, Inc.

Almond, D. and Edlund, L. (2007) Trivers–Willard at birth and one year: Evidence from US natality data 1983–2001. *Proceedings of the Royal Society B: Biological Sciences* 274(1624): 2491–2496.

Almond, D., Chay, K.Y. and Lee, D.S. (2005) The costs of low birth weight. *The Quarterly Journal of Economics* 120(3): 1031–1083.

Ananat, E.O., Gruber, J. and Levine, P. (2007) Abortion legalization and life-cycle fertility. *Journal of Human Resources* 42(2): 375–397.

Ananat, E.O., Gruber, J., Levine, P.B. and Staiger, D. (2009) Abortion and selection. *The Review of Economics and Statistics* 91(1): 124–136.

Ananat, E.O. and Hungerman, D.M. (2012) The power of the pill for the next generation: Oral contraception's effects on fertility, abortion, and maternal and child characteristics. *The Review of Economics and Statistics* 94(1): 37–51.

Angrist, J., Lavy, V. and Schlosser, A. (2010) Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics* 28(4): 773–824.

Angrist, J.D. and Evans, W.N. (1996) Schooling and labor market consequences of the 1970 state abortion reforms. NBER Working Papers 5406, National Bureau of Economic Research, Inc.

Angrist, J.D. and Evans, W.N. (1998) Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88(3): 450–477.

Bailey, M.J. (2006) More power to the pill: The impact of contraceptive freedom on women's life cycle labor supply. *The Quarterly Journal of Economics* 121(1): 289–320.

Bailey, M.J. (2009) Erratum and addendum. Posted at *The Quarterly Journal of Economics*.

Bailey, M.J. (2010) "Momma's got the pill": How Anthony Comstock and Griswold v. Connecticut shaped US childbearing. *American Economic Review* 100(1): 98–129.

Bailey, M.J. (2012) Reexamining the impact of family planning programs on US fertility: Evidence from the war on poverty and the early years of title X. *American Economic Journal: Applied Economics* 4(2): 62–97.

Bailey, M.J. (2013) Fifty years of family planning: New evidence on the long-run effects of increasing access to contraception. *Brookings Papers on Economic Activity* 46(1): 341–409.

Bailey, M.J., Guldi, M. and Hershbein, B.J. (2013) Further evidence on the internal validity of the early legal access research design. *Journal of Policy Analysis and Management* 32(4): 899–904.

Bailey, M.J., Hershbein, B. and Miller, A.R. (2012) The opt-in revolution? Contraception and the gender gap in wages. *American Economic Journal: Applied Economics* 4(3): 225–254.

Bailey, M.J., Malkova, O. and McLaren, Z.M. (2016) Does family planning increase children's opportunities? Evidence from the war on poverty and the early years of title X. Technical report, University of Michigan.

Barro, R.J. and Becker, G.S. (1989) Fertility choice in a model of economic growth. *Econometrica* 57(2): 481–501.

Becker, G.S. (1960) An economic analysis of fertility. In *Demographic and Economic Change in Developed Countries*, NBER chapters (pp. 209–240). National Bureau of Economic Research, Inc.

Becker, G.S. (1965) A theory of the allocation of time. *The Economic Journal* 75(299): 493–517.

Becker, G.S. and Barro, R.J. (1988) A reformulation of the economic theory of fertility. *The Quarterly Journal of Economics* 103(1): 1–25.

Becker, G.S. and Lewis, H.G. (1973) On the interaction between the quantity and quality of children. *Journal of Political Economy* 81(2): S279–288.

Becker, G.S. and Tomes, N. (1976) Child endowments and the quantity and quality of children. *Journal of Political Economy* 84(4): S143–162.

Becker, G.S. and Tomes, N. (1986) Human capital and the rise and fall of families. *Journal of Labor Economics* 4(3): S1–S39.

Becker, G.S., Murphy, K.M. and Tamura, R. (1990) Human capital, fertility, and economic growth. *Journal of Political Economy* 98(5): S12–37.

Becker, S.O., Cinnirella, F. and Woessmann, L. (2010) The trade-off between fertility and education: Evidence from before the demographic transition. *Journal of Economic Growth* 15(3): 177–204.

Behrman, J., Taubman, P. and Pollak, R. (1982) Parental preferences and provision for progeny. *Journal of Political Economy* 90(1): 52–73.

Ben-Porath, Y. (1976) Fertility response to child mortality: Micro data from Israel. *Journal of Political Economy* 84(4): S163–178.

Ben-Porath, Y. and Welch, F. (1972) Chance, child traits, and choice of family size. Report R-1117-NIH/RF, RAND Corporation.

Bentancor, A. and Clarke, D. (Forthcoming) Assessing plan B: The effect of the morning after pill on women and children. *The Economic Journal*.

Bhalotra, S. and Clarke, D. (2015) The twin instrument and the estimation of the quality-quantity trade-off. Mimeo, University of Essex and University of Oxford.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2005) The more the merrier? The effect of family size and birth order on children's education. *The Quarterly Journal of Economics* 120(2): 669–700.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2008) Staying in the classroom and out of the maternity ward? The effect of compulsory schooling laws on teenage births. *The Economic Journal* 118(530): 1025–1054.

Black, S.E., Devereux, P.J. and Salvanes, K.G. (2010) Small family, smart family? Family size and the iq scores of young men. *Journal of Human Resources* 45(1): 33–58.

Bleakley, H. and Lange, F. (2009) Chronic disease burden and the interaction of education, fertility, and growth. *The Review of Economics and Statistics* 91(1): 52–65.

Bongaarts, J. and Sinding, S. (2011) Population policy in transition in the developing world. *Science* 333(6042): 574–576.

Bougma, M., LeGrand, T. and Kobiané, J.-F. Fertility decline and child schooling in urban settings of Burkina Faso. *Demography* 52(1): 281–313.

Brinch, C.N., Mogstad, M. and Wiswall, M. (2012) Beyond LATE with a discrete instrument. Heterogeneity in the quantity-quality interaction of children. Discussion Papers 703, Statistics Norway, Research Department.

Bronars, S.G. and Grogger, J. (1994) The economic consequences of unwed motherhood: Using twin births as a natural experiment. *The American Economic Review* 84(5): 1141–1156.

Butcher, K.F. and Case, A. (1994) The effect of sibling sex composition on women's education and earnings. *The Quarterly Journal of Economics* 109(3): 531–563.

Cáceres-Delpiano, J. (2006) The impacts of family size on investment in child quality. *Journal of Human Resources* 41(4): 738–754.

Cáceres-Delpiano, J. (2008) Keeping the best for last. Impact of fertility on mother's employment. Evidence from developing countries. Economics Working Papers we086832, Universidad Carlos III, Departamento de Economía.

Cáceres-Delpiano, J. (2012) Can we still learn something from the relationship between fertility and mother's employment? Evidence from developing countries. *Demography* 49(1): 151–174.

Charles, K.K. and Melvin, S., Jr, (2006) Abortion legalization and adolescent substance use. *Journal of Law and Economics* 49(2): 481–505.

Chevalier, A. and Viitanen, T.K. (2003) The long-run labour market consequences of teenage motherhood in Britain. *Journal of Population Economics* 16(2): 323–343.

Christensen, F. (2012) The pill and partnerships: The impact of the birth control pill on cohabitation. *Journal of Population Economics* 25(1): 29–52.

Chun, H. and Oh, J. (2002) An instrumental variable estimate of the effect of fertility on the labour force participation of married women. *Applied Economics Letters* 9(10): 631–634.

CIA., (2014) *The World Fact Book*. Washington DC: Central Intelligence Agency.

Conley, D. and Glauber, R. (2006) Parental educational investment and children's academic risk: Estimates of the impact of sibship size and birth order from exogenous variation in fertility. *The Journal of Human Resources* 41(4): 722–737.

Cook, P.J., Parnell, A.M., Moore, M.J. and Pagnini, D. (1999) The effects of short-term variation in abortion funding on pregnancy outcomes. *Journal of Health Economics* 18(2): 241–257.

Cristia, J.P. (2008) The effect of a first child on female labor supply: Evidence from women seeking fertility services. *The Journal of Human Resources* 43(3): 487–510.

Currie, J., Nixon, L. and Cole, N. (1996) Restrictions on Medicaid funding of abortion: Effects on birthweight and pregnancy resolutions. *The Journal of Human Resources* 31(1): 159–187.

Dahl, G.B. and Moretti, E. (2008) The demand for sons. *The Review of Economic Studies* 75(4): 1085–1120.

Dang, H.-A. and Rogers, H. (2013) The decision to invest in child quality over quantity: Household size and household investment in education in Vietnam. Policy Research Working Paper Series 6487, The World Bank.

Darroch, J.E. (2013) Trends in contraceptive use. *Contraception* 87(3): 259–263.

Dayioğlu, M., Kirdar, M.G. and Tansel, A. (2009) Impact of sibship size, birth order and sex composition on school enrolment in urban Turkey. *Oxford Bulletin of Economics and Statistics* 71(3): 399–426.

de la Croix, D. and Doepke, M. (2003) Inequality and growth: Why differential fertility matters. *American Economic Review* 93(4): 1091–1113.

De Tray, D.N. (1973) Child quality and the demand for children. *Journal of Political Economy* 81(2): S70–95.

Deheija, R., Pop-Eleches, C. and Samii, C. (2015) From local to global: External validity in a fertility natural experiment. Mimeo, Available for download at http://users.nber.org/˜rdehejia/papers/dehejia_pop-eleches_samii.pdf, (accessed on April 12, 2016).

Desai, S. (1995) When are children from large families disadvantaged? Evidence from cross-national analyses. *Population Studies* 49(2): 195–210.

DiCenso, A., Guyatt, G., Willan, A. and Griffith, L. (2002) Interventions to reduce unintended pregnancies among adolescents: Systematic review of randomised controlled trials. *British Medical Journal* 324(7351): 1426–1434.

Doepke, M. (2015) Gary Becker on the quantity and quality of children. *Journal of Demographic Economics* 81(1): 59–66.

Duflo, E. (2001) Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91(4): 795–813.

Durrance, C.P. (2013) The effects of increased access to emergency contraception on sexually transmitted disease and abortion rates. *Economic Inquiry* 51(3): 1282–1695.

Easterlin, R.A. (1975) An economic framework for fertility analysis. *Studies in Family Planning* 6(3): 54–63.

Ebenstein, A. (2009) When is the local average treatment close to the average?: Evidence from fertility and labor supply. *Journal of Human Resources* 44(4): 955–975.

Edlund, L. and Machado, C. (2015) How the other half lived: Marriage and emancipation in the age of the pill. *European Economic Review* 80(C): 295–309.

Enke, S. (1966) The economic aspects of slowing population growth. *The Economic Journal* 76(301): 44–56.

Enke, S. (1971) Economic consequences of rapid population growth. *The Economic Journal* 81(324): 800–811.

Fitzsimons, E. and Malde, B. (2010) Empirically probing the quantity-quality model. IFS Working Papers W10/20, Institute for Fiscal Studies.

Fitzsimons, E. and Malde, B. (2014) Empirically probing the quantity-quality model. *Journal of Population Economics* 27(1): 33–68.

Fleisher, B.M., Rhodes, J. and George, F. (1979) Fertility, women's wage rates, and labor supply. *The American Economic Review* 69(1): 14–24.

Fletcher, J. (2012) The effects of teenage childbearing on the short- and long-term health behaviors of mothers. *Journal of Population Economics* 25(1): 201–218.

Fletcher, J.M. and Wolfe, B.L. (2009) Education and labor market consequences of teenage childbearing: Evidence using the timing of pregnancy outcomes and community fixed effects. *Journal of Human Resources* 44(2): 303–325.

Francesconi, M. (2002) A joint dynamic model of fertility and work of married women. *Journal of Labour Economics* 20(2): 336–380.

Geronimus, A.T. and Korenman, S. (1992) The socioeconomic consequences of teen childbearing reconsidered. *The Quarterly Journal of Economics* 107(4): 1187–1214.

Goldin, C. (2006) The quiet revolution that transformed women's employment, education, and family. Working paper no. 11953, NBER.

Goldin, C. and Katz, L.F. (2002) The power of the pill: Oral contraceptives and women's career and marriage decisions. *Journal of Political Economy* 110(4): 730–770.

Greenwood, J., Guner, N. and Vandenbroucke, G. (2016) Family economics writ large. Working Papers 2016-26, Federal Reserve Bank of St. Louis.

Gronau, R. (1973) The effect of children on the housewife's value of time. *Journal of Political Economy* 81(2): S168–S199.

Gross, T., Lafortune, J. and Low, C. (2014) What happens the morning after? The costs and benefits of expanding access to emergency contraception. *Journal of Policy Analysis and Management* 33(1): 70–93.

Gruber, J., Levine, P. and Staiger, D. (1999) Abortion legalization and child living circumstances: Who is the "marginal child"? *The Quarterly Journal of Economics* 114(1): 263–291.

Guldi, M. (2008) Fertility effects of abortion and birth control pill access for minors. *Demography* 45(4): 17–827.

Haavelmo, T. (1943) The statistical implications of a system of simultaneous equations. *Econometrica* 11(1): 1–12.

Haavelmo, T. (1944) The probability approach in econometrics. *Econometrica* 12(Supplement): iii–vi+1–115.

Hanushek, E.A. (1992) The trade-off between child quantity and quality. *Journal of Political Economy* 100(1): 84–117.

Heckman, J. and Pinto, R. (2015) Causal analysis after Haavelmo. *Econometric Theory* 31: 115–151.

Heckman, J.J. and Willis, R.J. (1976) Estimation of a stochastic model of reproduction: An econometric approach. In *Household Production and Consumption*, NBER chapters (pp. 99–146). Boston, USA: National Bureau of Economic Research, Inc.

Holmlund, H. (2005) Estimating long-term consequences of teenage childbearing: An examination of the siblings approach. *The Journal of Human Resources* 40(3): 716–743.

Hotz, V.J., McElroy, S.W. and Sanders, S.G. (2005) Teenage childbearing and its life cycle consequences: Exploiting a natural experiment. *The Journal of Human Resources* 40(3): 683–715.

Hotz, V.J. and Miller, R.A. (1988) An empirical analysis of life cycle fertility and female labor supply. *Econometrica* 56(1): 91–118.

Hotz, V.J., Mullin, C.H. and Sanders, S.G. (1997) Bounding causal effects using data from a contaminated natural experiment: Analysing the effects of teenage childbearing. *Review of Economic Studies* 64(4): 575–603.

Imbens, G.W. and Angrist, J.D. (1994) Identification and estimation of local average treatment effects. *Econometrica* 62(2): 467–475.

Jacobsen, J.P., Pearce III, J.W. and Rosenbloom, J.L. (1999) The effects of childbearing on married women's labor supply and earnings: Using twin births as a natural experiment. *Journal of Human Resources* 34(3): 449–474.

Joshi, S. and Schultz, T. (2013) Family planning and women's and children's health: Long-term consequences of an outreach program in Matlab, Bangladesh. *Demography* 50(1): 149–180.

Joyce, T. (2013) If only policy analysis were so easy. *Journal of Policy Analysis and Management* 32(4): 897–898. ISSN 1520-6688.

Kane, T. and Staiger, D. (1996) Teen motherhood and abortion access. *Quarterly Journal of Economics* 111(2): 467–506.

Keane, M.P. (2010) Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics* 156(1): 3–20.

Kearney, M.S. and Levine, P.B. (2009) Subsidized contraception, fertility, and sexual behavior. *The Review of Economics and Statistics* 91(1): 137–151.

Kearney, M.S. and Levine, P.B. (2012) Why is the teen birth rate in the United States so high and why does it matter? *The Journal of Economic Perspectives* 26(2): 141–166.

Kim, J. and Aassve, A. (2006) Fertility and its consequence on family labour supply. IZA Discussion Papers 2162, Institute for the Study of Labor (IZA).

Koopmans, T.C. (1947) Measurement without theory. *The Review of Economics and Statistics* 29(3): 161–172.

Kumar, S. and Kugler, A. (2011) Testing the children quantity-quality trade-off in India. MPRA Paper 42487, University Library of Munich, Germany.

Lee, J. (2008) Sibling size and investment in children's education: An Asian instrument. *Journal of Population Economics* 21(4): 855–875.

Levine, D.I. and Painter, G. (2003) The schooling costs of teenage out-of-wedlock childbearing: Analysis with a within-school propensity-score-matching estimator. *The Review of Economics and Statistics* 85(4): 884–900.

Levine, P., Staiger, D., Kane, T. and Zimmerman, D. (1999) *Roe v Wade* and American fertility. *American Journal of Public Health* 89(2): 199–203.

Levine, P., Trainor, A. and Zimmerman, D. (1996a) The effect of Medicaid abortion funding restrictions on abortions, pregnancies and births. *Journal of Health Economics* 15(5): 555–578.

Levine, P.B., Trainor, A.B. and Zimmerman, D.J. (1996b) The effect of Medicaid abortion funding restrictions on abortions, pregnancies and births. *Journal of Health Economics* 15(5): 555–578.

Li, H., Zhang, J. and Zhu, Y. (2008) The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography* 45: 223–243.

Lundborg, P., Plug, E. and Rasmussen, A.W. (2014) Fertility effects on female labor supply: IV evidence from IVF treatments. IZA Discussion Papers 8609, Institute for the Study of Labor (IZA).

Maralani, V. (2008) The changing relationship between family size and educational attainment over the course of socioeconomic development: Evidence from Indonesia. *Demography* 45(3): 693–717.

Marteleto, L.J. and de Souza, L.R. (2012) The changing impact of family size on adolescents' schooling: Assessing the exogenous variation in fertility using twins in Brazil. *Demography* 49(4): 1453–1477.

Miller, A. (2011) The effects of motherhood timing on career path. *Journal of Population Economics* 24(3): 1071–1100.

Miller, A.R. (2009) Motherhood delay and the human capital of the next generation. *American Economic Review* 99(2): 154–158.

Miller, G. (2010) Contraception as development? New evidence from family planning in Colombia. *The Economic Journal* 120(545): 709–736.

Millimet, D.L. and Wang, L. (2011) Is the quantity-quality trade-off a trade-off for all, none, or some? *Economic Development and Cultural Change* 60(1): 155–195.

Moffitt, R. (2005) Remarks on the analysis of causal relationships in population research. *Demography* 42(1): 91–108.

Mogstad, M. and Wiswall, M. (2016) Testing the quantity–quality model of fertility: Estimation using unrestricted family size models. *Quantitative Economics* 7(1): 157–192.

Molyneux, J.W. and Gertler, P.J. (2000) The impact of targeted family planning programs in Indonesia. *Population and Development Review* 26: 61–85.

Myrskyl, M.ä, Kohler, H.-P. and Billari, F.C. (2009) Advances in development reverse fertility declines. *Nature* 460(7256): 741–743.

O'Dowd, M.J. and Philipp, E.E. (1994) *The History of Obstetrics and Gynaecology*. New York: Parthenon.

Ponczek, V. and Souza, A.P. (2012) New evidence of the causal effect of family size on child quality in a developing country. *Journal of Human Resources* 47(1): 64–106.

Pop-Eleches, C. (2006) The impact of an abortion ban on socioeconomic outcomes of children: Evidence from Romania. *Journal of Political Economy* 114(4): 744–773.

Pop-Eleches, C. (2010) The supply of birth control methods, education and fertility: Evidence from Romania. *Journal of Human Resources* 45(4): 971–997.

Qian, N. (2009) Quantity-quality and the one child policy: The only-child disadvantage in school enrollment in rural China. NBER Working Papers 14973, National Bureau of Economic Research, Inc.

Ribar, D.C. (1994) Teenage fertility and high school completion. *The Review of Economics and Statistics* 76(3): 413–424.

Ribar, D.C. (1999) The socioeconomic consequences of young women's childbearing: Reconciling disparate evidence. *Journal of Population Economics* 12(4): 547–565.

Rosenzweig, M.R. and Schultz, T.P. (1985) The demand for and supply of births: Fertility and its life cycle consequences. *The American Economic Review* 75(5): 992–1015.

Rosenzweig, M.R. and Schultz, P.T. (1987) Fertility and investments in human capital: Estimates of the consequence of imperfect fertility control in Malaysia. *Journal of Econometrics* 36 (1–2): 163–184.

Rosenzweig, M.R. and Wolpin, K.I. (1980a) Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica* 48(1): 227–240.

Rosenzweig, M.R. and Wolpin, K.I. (1980b) Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy* 88 (2): 328–348.

Rosenzweig, M.R. and Wolpin, K.I. (1986) Evaluating the effects of optimally distributed public programs: Child health and family planning interventions. *The American Economic Review* 76(3): 470–482.

Rosenzweig, M.R. and Zhang, J. (2009) Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy. *Review of Economic Studies* 76(3): 1149–1174.

Sanhueza, C. (2009) Tamaño de familia y orden de nacimiento en Chile: Usando gemelos como experimento natural. *Revista de Analisis Económico - Economic Analysis Review* 24(2): 133–152.

Schultz, P.T. (2008) Population policies, fertility, women's human capital, and child quality. In P. Schultz and J. Strauss (eds.), *Handbook of Development Economics*, Vol. 4 (pp. 3249–3303). Amsterdam: Elsevier Science B.V.

The World Bank. (2015) Data bank. http://data.worldbank.org, 2015. Consulted 31 May 2015.

Todd, P.E. and Wolpin, K.I. (2006) Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review* 96(5): 1384–1417.

Trivers, R.L. and Willard, D.E. (1973) Natural selection of parental ability to vary the sex ratio of offspring. *Science* 179(4068): 90–92.

United Nations. (2014). World fertility report 2013: Fertility at the extremes. Technical report, Department of Economic and Social Affairs: Population Division.

United Nations High Commissioner for Human Rights. (1968) Proclamation of Teheran: Proclaimed by the International Conference on Human Rights at Teheran on 13 May. URL http://web.archive.org/web/20071017025912/http://www.unhchr.ch/html/menu3/b/b_tehern.htm. [Online; accessed 30 August 2015].

Willis, R.J. (1973) A new approach to the economic theory of fertility behavior. *Journal of Political Economy* 81(2): S14–S64.

Wolpin, K. (2013) *The Limits of Inference without Theory*, Vol. 1 (1st ed.). Cambridge MA: The MIT Press.